DOCUMENT RESUME

ED 222 523 TM 820 670

AUTHOR Ludlow, Larry H.
TITLE A Residual Is More Than a Chi-Square: Patterns in Attitudes Toward Blindness.
PUB DATE Mar 82
NOTE 21p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New York, NY, March 20-22, 1982); Some figures are marginally legible due to small print.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Attitude Measures; *Behavior Patterns; *Blindness; Caseworkers; Group Behavior; Item Analysis; Psychometrics; Rating Scales; Rehabilitation Programs; *Statistical Analysis; *Test Construction
IDENTIFIERS *Attitude Toward Blindness Questionnaire; Residuals (Statistics); Veterans Administration

ABSTRACT

An Attitude Toward Blindness Questionnaire (ATBQ) was developed to assess individual progress through a Veterans Administration blind rehabilitation program. The instrument is also meant to measure attitudes toward blindness for blind persons, rehabilitation workers, and a "naive" group without contact with the blind. A method of assessing model fit by a Rasch Rating Scale Model was used to make group comparisons for any chosen demographic stratification without re-calibration for every set of comparisons. A conventional analysis of group and item distribution was undertaken, with an alternative analysis further recognizing that group differences must also exist within the residuals from the original overall analysis. The residuals resulted from the difference between observed and modeled responses, and were scaled to yield a standardized residual. The analysis calibrated items to uncover peculiar individual response patterns from a combined sample of workers and blind persons, facilitating the identification of specific rehabilitation needs. Estimate plot, matrix techniques and graphs for comparing group residuals, and the uses of negative residual distributions in the calibrations are described. (CM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED222523

A Residual is More Than a Chi-Square:

Patterns in Attitudes Toward Blindness

Larry H. Ludlow

University of Chicago

and

The Rehabilitative Engineering Research
and Development Laboratory
Hines Veterans Administration Hospital
Hines, Illinois

Paper Presented at:

The National Council on Measurement in Education
New York, March 1982

TM 820 670

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

L. H. Ludlow

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The Blind Rehabilitation Center at USVA Hospital, Hines, Illinois is interested in the development of an instrument that will measure attitudes toward blindness. One purpose of this instrument will be the assessment of an individual's progress through the rehabilitation program. Not only must the instrument measure what is intended but it must measure the same thing for blind persons, rehabilitation workers and persons without any special contact with the blind. This last group will be referred to as the "naive" group.

The Center calls their instrument the Attitude Toward Blindness Questionnaire (ATBQ). It is composed of four distinct sub-scales. This paper discusses one way of assessing model fit by looking at the residuals remaining after one ATBQ sub-scale has been fit by the Rasch psychometric model referred to as the Rating Scale Model (Andrich, 1978a, 1978b; Masters, 1980; Wright and Masters, 1982). Specifically, this paper does not discuss how residuals may be squared and summed to form approximate chi-squares as is typically done. Rather, the techniques considered involve a variety of plots, graphs and tables.

Figure 1 lists the items in the Can Negative sub-scale. The scoring pattern ranges from 0: strongly agree to 3: strongly disagree. A high response reflects a positive attitude because of disagreement with the item. A brief consideration of the items shows how this scale works. Namely, these items ask questions about things which the blind might do but if they were to do it the consequence would be negative. Our first question then is whether or not these items may be placed

along a continuum in such a way that at one end are items that most people would agree with and at the other end are items that most people would disagree with.

The essential results of the analysis are shown on Figure 2. There were 146 people and 13 items. To the left of the central vertical line are the people represented by x's. The items are to the right of the line. The positioning of the items and people defines the meaning of the Can Negative sub-scale. At the top of the distributions are people with high positive attitudes and items that most people tend to agree with. At the bottom are people with negative attitudes and items which most people tend to disagree with. Only the people located above I208 are expected to disagree that the blind make the sighted feel uneasy. Likewise, the person at the bottom of the scale is the one with the highest probability of agreeing that blindness is a form of punishment for one's sins. The ends of the distribution seem sensible. The arrangement of the items within the distribution is also reasonable. In the right margin of the page are three clusters of items. The upper cluster of items seem related to the social effect of blindness. The middle cluster seems related to the practical effect of blindness. The lower cluster of items tap into increasing more severe detrimental effects of blindness. For our purposes this set of items seems to measure one aspect of the attitude one could hold toward blind persons.

The next question is whether this scale remains the same for blind persons and rehabilitation workers. In the original analysis of these

data the "naive" group was also included. Their results are not included here because they consistently fell between the patterns specific to the blind and workers. Hence their discussion adds nothing to the purpose of presenting a useful methodolgy.

One thing we could do would be to plot the attitude measures for the workers against the measures for the blind. This would tell us of the relatively more or less positive attitude one group or the other held toward these items but it would not tell us if the item locations remained invariant. The issue of invariance is the more important one because it tests whether the items mean the same thing for the groups. The group attitude measures were plotted against one another and it was seen that the workers as a group had a noticeably more positive attitude toward these items than did the blind.

Figure 3 shows an analysis that is commonly done as a test of invariance. The workers are analyzed separately as are the blind. Their respective item calibrations are plotted against one another and a confidence interval is constructed. Two items fall outside the 2 standard error confidence interval constructed for these data. In the upper left quadrant is item I106 (blind employees require supervisors with special training). This item was harder for the blind to disagree with than it was for the workers. In the lower right quadrant is item I101 (blind people use their blindness as an excuse not to work). This item was harder for the workers to disagree with than it was for the blind. Obviously, this information should be communicated to the workers.

The conventional analysis usually stops at this point. There is, however, another way to see the same information plus additional item and person detail. The alternative analysis is the result of recognizing that group differences seen in Figure 3 must also exist within the residuals from the original calibration with everyone combined in the overall analysis. This fact enables us to make group comparisons for any demographic stratification we choose to make without re-calibrating our instrument for every set of comparisons.

The residuals to be discussed result from the difference between the observed response and the modelled response. This quotient is then divided be a scaling factor to yield a standardized residual. This form of residual is chosen because it is in a familiar metric.

One general way to look at residuals is to plot them against the model estimates. For instance, we can plot them against the attitude measures and then plot them against the item calibrations. These plots are frequently interesting. It is common to find a diagonal pattern in this type of plot because as the estimate becomes more extreme, surprising responses tend to occur. What is of interest, then, is not so much that a diagonal pattern occured but whether or not the pattern was there for all the groups of interest when the residuals for each group are plotted.

Another general way is to build a matrix highlighting patterns of large residuals. A criteria like "any residual beyond plus or minus 2

standardized units" could be defined. The matrix then would contain rows of people and their residuals from the specific items. The items could be arranged as the columns. This matrix is useful for comparing patterns between two or more groups. The residual by estimate plot and the matrix technique are illustrated in a previous paper and are not presented in this analysis (Ludlow, 1981).

Before we directly address the group difference question it is of some interest to inquire into the type of distribution the residuals form. One way to inspect their distribution is shown in Figure 4. The technique is illustrated by two items. The left item represents the typical pattern found in these data. The right item illustrates the most peculiar pattern. The two top pictures are referred to as hanging rootograms. They illustrate the fit of the observed frequencies to that expected for a normal distribution. The ordinate is the root of the normal density function evaluated at the bin midpoints. The spikes coming down from the curve are of a length equal to the root of the number in the bin divided by the product of the sample size and bin width. Spikes ending between the two dashed lines represent an acceptable fit to the normal distribution within plus or minus two standard error units. A utility of this type of frequency plot is the use of a constant standard error for the entire distribution. What we see for both items is an abundance of negative residuals and a truncated right tail.

This observation is confirmed by the statistics printed in the middle of the page. G1 (skewness) is negative, hence a negative skew and

G2 (kurtosis) is positive, hence a bunching up in the middle. The bottom plots are examples of rankit plots. The vertical axis is the expected normal order statistic distribution and the horizontal axis is the observed order statistic distribution. The observed order statistics are simply the residuals arranged from negative to positive. We see again more in the left tail and less in the right tail than expected. The point of this information is that there are quite a few surprising agree responses.

We must ask how these large negative residuals are coming about. Are they because people actually are responding with agree or strongly agree or is it simply because people have such high attitude measures that even a disagree response rather than strongly disagree can result in a large negative residual. Further we need to check whether there is a pattern to who made the surprising responses.

Figure 5 shows an interesting way to address both questions simultaneously. I242 was chosen because it had the most peculiar pattern and seemed a reasonable choice as an example for investigating fit. These plots are of the residuals plotted against the attitude measures. What we find is a very interesting pattern. This pattern results because we are modelling discrete responses. It is easy enough to show that if all the difficulties were the same value and all the measures were the same value then only 4 residuals can result because only four responses are possible. Thus as the estimates spread out the residuals spread out but form a characteristic pattern for each of the observed responses. In

particular as the measures increase the residuals from the most positive responses (here the SD response) will aysmptotically approach zero. Likewise, as the measures decrease the residuals from the most negative responses aysmptotically approach zero. Again, since we know the expected pattern for this type of plot the interesting information is that seen when the plot does not take that form.

The upper plot is for the combined worker and blind residuals. The lower left plot is for the blind only and the lower right plot is for the workers. There is an immediately apparent difference between the two bottom plots. Three features are important. One is that negative residuals to this item are indeed being made by some agree and strongly agree responses. Two, these responses are being made by the blind. Three, there is a very peculiar rehabilitation worker. This information is of obvious importance to the rehabilitation process. Before leaving this type of plot I point out that when a person is flagged like the one circled in the worker only plot it is possible to build an individual picture for a person by plotting the person's residuals against his observed responses. This picture shows the surprise associated with each of his responses.

If we believe means and standard deviations are useful first approximations for residual distributions, we can summarize group differences in graphs like those in Figure 6. In this type of graph we simply count the number of residuals that fall at any point along the line and write that number above the line. For the first item on this

page (I101 - work excuse) we see that the workers tend to be negatively skewed while the blind are fairly evenly distributed around zero. This simply means there are more negative residuals (surprising agrees) than we expected from people of their attitude position. Looking at the next item we see the opposite pattern. Here the workers are characterized by a positive pattern (surprising disagree) of residuals. The importance of these graphs is that they are the same two items flagged in Figure 3 where the separate item calibrations are plotted. The generalization we may draw is that there is a direct relation between the mean residual on an item for a particular group and the separate group item calibration. Simply put, when the group item mean residual is negative, the separate item calibation is greater than that derived from the overall calibration. Their overall analysis expected scores are too high for them. Conversely, when the group item mean residual is positive, the separate item calibration is less than that derived from the overall calibration. Their overall analysis expected scores are too low for them.

The bottom two items are included because they are the next set of items with the largest group mean differences. These item were not flagged in Figure 3 but they too have useful diagnostic information. Finally, we can build a summary of these mean relations by plotting the group item residual means against one another as in Figure 7.

In this plot the blind are plotted on the vertical axis, the workers on the horizontal. The interesting quadrants are the upper left and lower right. In the upper left the workers tend to agree with the items

while the blind tend to disagree with them. These would seem to be realities the workers recognize but that the blind reject. In the lower right the workers tend to disagree while the blind tend to agree. These would seem to be areas of fears and worries specific to blindnedss that the workers are not as aware of as they might be. For the purpose of rehabilitation this plot may be the most diagnostic of all the previous ones. For this plot to be useful it is essential that standard deviations be near one for both groups on all items because the mean can be seriously distorted. There are items in the 1st and 3rd quadrants because of the large number of "naive" persons included in the overall calibration.

In conclusion one might think it to be strange logic to calibrate items from a combined sample of workers and blind persons when it is a prior suspected that they might differ in their attitudes. For one to hold that attitude toward analysis would be to miss important pieces of information because the plot in Figure 3 does not convey all the information possible. The utility of this detailed residual analysis was demonstrated by the uncovering of individuals from both groups who had peculiar response patterns. Their peculiarities suggest a specific rehabilitation program tailored to their needs. This diagnosis does not in and of itself constitute a rejection of the model. Without a model having first been specified we would not have found the patterns we did. Finally , for those who would tend to rely primarily on statistical analyses we agree that most of these patterns would not meet statistically significant standards and, typically, would not recieve much attention.

For the purpose of rehabilitation, however, strict statistical standards might not always be the most appropriate to follow. That was the case here because many of the peculiarities uncovered would not have otherwise been seen and could not have been incoporated into rehabilitation programs, which, in the final analysis, is most crucial at the individual level.

References

Andrich, D. A rating formulation for ordered response categories. Psychometrika, 1978a, 43, 561-573.

Andrich, D. Scaling attitude items constructed and scored in the Likert tradition. Educational and Psychological Measurement, 1978b, 38, 665-680.

Ludlow, L.H. An exploratory investigation of Rasch model residuals. American Educational Research Association meeting, Los Angeles, April 1981, ERIC Document: ED 206 652.

Masters, G.N. A Rasch model for rating scales. Doctoral dissertation, University of Chicago, 1980.

Wright, B.D. and Masters, G.N. Rating Scale Analysis. Chicago: MESA Press, 1982.

Figure 1

Attitude Toward Blindness Questionnaire:
Can Negative Items - Form 1

Scoring Pattern:

0: Strongly Agree 1: Agree 2: Disagree 3: Strongly Disagree

Item Name	Full Description (Brief label)
I101	Blind people use their blindness as an excuse not to work. (work excuse)
I103	Blind employees can make their co-workers uneasy. (co-workers uneasy)
I104	Blind people have trouble working for large companies. (large companies)
I106	Blind employees require supervisors with special training. (supervisors)
I108	Blind people add extra costs to an employers labor bill. (extra costs)
I110	Blind people tend to daydream on the job. (daydream)
I202	Blind people are uneasy in the presence of others. (uneasy others)
I206	Blind people make sighted people feel uneasy. (sighted uneasy)
I208	Blind people feel more social pain than sighted people. (social pain)
I230	A blind person can become mentally ill more easily than a sighted person. (mentally ill)
I232	A blind person can easily put a lot of stress on family bonds. (family stress)
I234	Blind people are a financial burden to their family. (financial burden)
I242	Blindness is a form of punishment for one's sins. (sin)

Figure 2

SCALE MAP SHOWING POSITIONS OF PEOPLE AND ITEMS ON THE VARIABLE

SCORE (FREQ)	PERSON POSITION	PEOPLE(N=148)	ITEMS(L= 13)	VALUE	ITEM DESCRIPTION	
38(1)	5.02	X				
37(1)	4.28	X				
36(3)	3.79	XXX				
35(1)	3.42	X				
34(5)	3.10	XXXXX				
33(0)	2.83					
32(1)	2.58	X				
31(6)	2.34	XXXXXX				
30(10)	2.12	XXXXXXXXXX				
29(5)	1.91	XXXXX				
28(5)	1.71	XXXXX				
27(8)	1.52	XXXXXXXX				
26(5)	1.32	XXXXX	I206	1.40	SIGHTED UNEASY	Social effect
			I103 I232	1.31	CO-WORKERS UNEASY FAMILY STRESS	Social effect
25(15)	1.13	XXXXXXXXXXXXXX				
24(12)	0.95	XXXXXXXXXXXX				
23(12)	0.76	XXXXXXXXXXXX				
22(16)	0.57	XXXXXXXXXXXXXXX	I208	0.64	SOCIAL PAIN	Practical effect
21(13)	0.38	XXXXXXXXXXXXX	I106	0.36	SUPERVISORS	Practical effect
20(5)	0.19	XXXXX				
19(9)	-0.00	XXXXXXXXX	I108 I234	-0.01	EXTRA COSTS FINANICIAL	Practical effect
18(8)	-0.20	XXXXXXXX	I104	-0.24	COMPANIES	Practical effect
17(2)	-0.39	XX	I202 I230	-0.41	UNEASY OTHERS MENTALLY ILL	Practical effect
16(1)	-0.59	X				Detrimental effect
15(0)	-0.80					Detrimental effect
14(0)	-1.01		I110	-0.95	DAYDREAM	Detrimental effect
			I101	-0.96	WORK EXCUSE	Detrimental effect
13(1)	-1.22	X				Detrimental effect
12(0)	-1.44					Detrimental effect
11(1)	-1.65	X				Detrimental effect
10(0)	-1.89					Detrimental effect
9(0)	-2.13		I242	-2.03	SIN PUNISHMENT	Detrimental effect

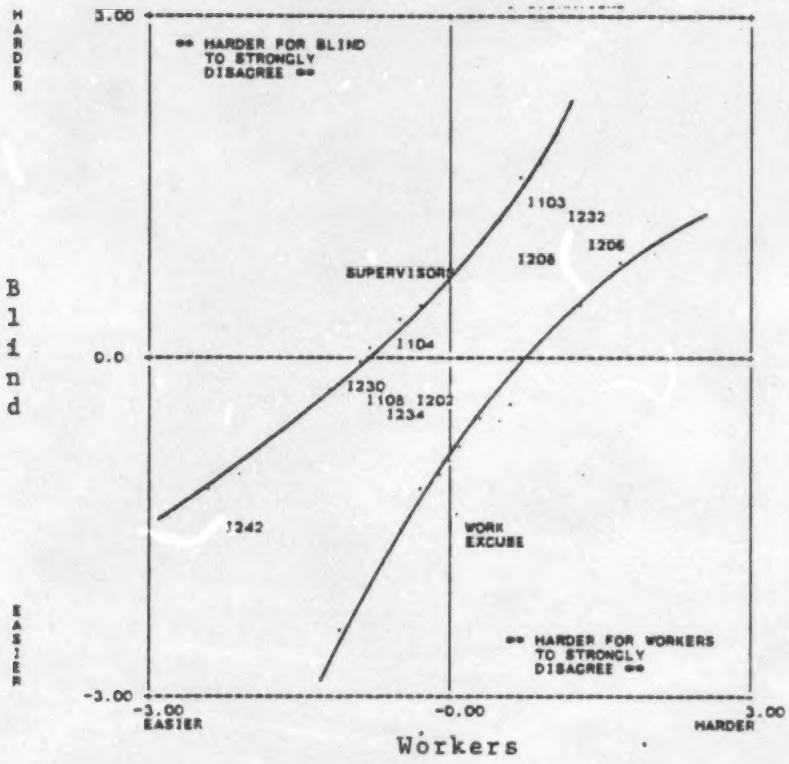
PERSONS WITH HIGH RAW SCORES-HIGH POSITIVE ATTITUDE

ITEMS WITH LOW RAW SCORES-DIFFICULT TO STRONGLY DISAGREE WITH

PERSONS WITH LOW RAW SCORES-LOW POSITIVE ATTITUDE

ITEMS WITH HIGH RAW SCORES-NOT TOO DIFFICULT TO STRONGLY DISAGREE WITH

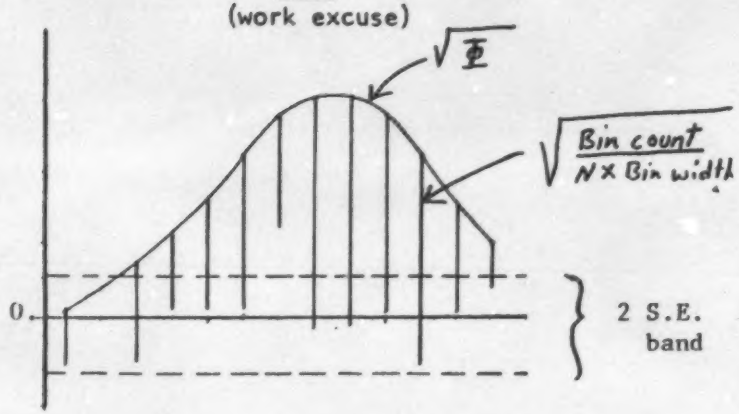
Worker item difficulties
plotted against
Blind item difficulties



THE CONFIDENCE INTERVAL REPRESENTS 2. STANDARD ERRORS

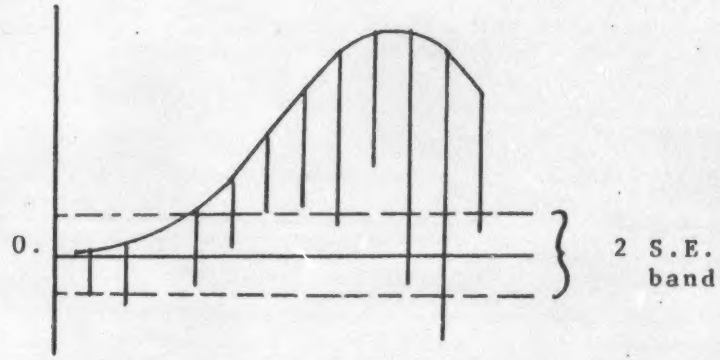
Hanging Rootograms

I101
(work excuse)



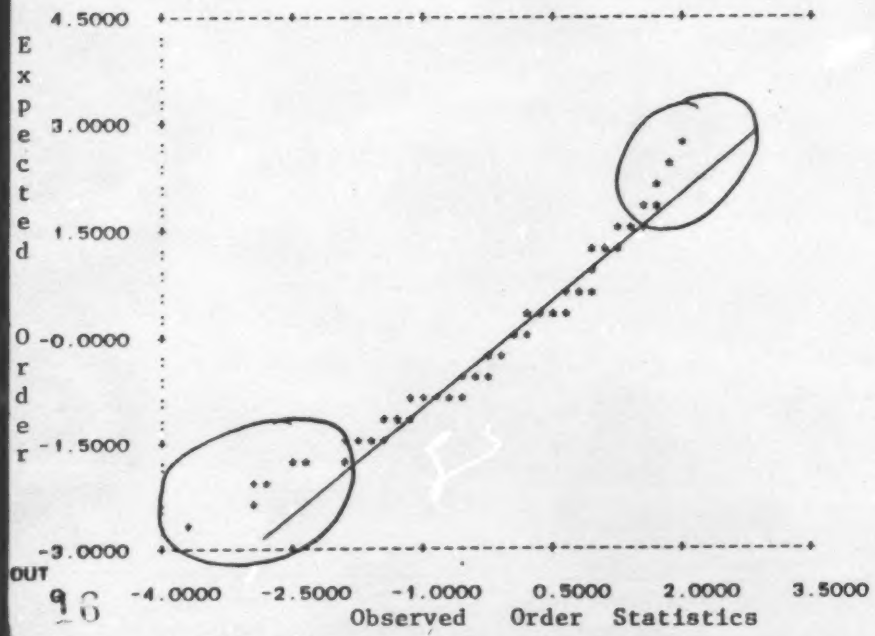
$\bar{X}$= -.018 S^2= 1.14 G1= -.77 G2= .8

I242
(sin)



$\bar{X}$= .023 S^2= 1.06 G1= -1.79 G2= 3.5

Rankit Plots



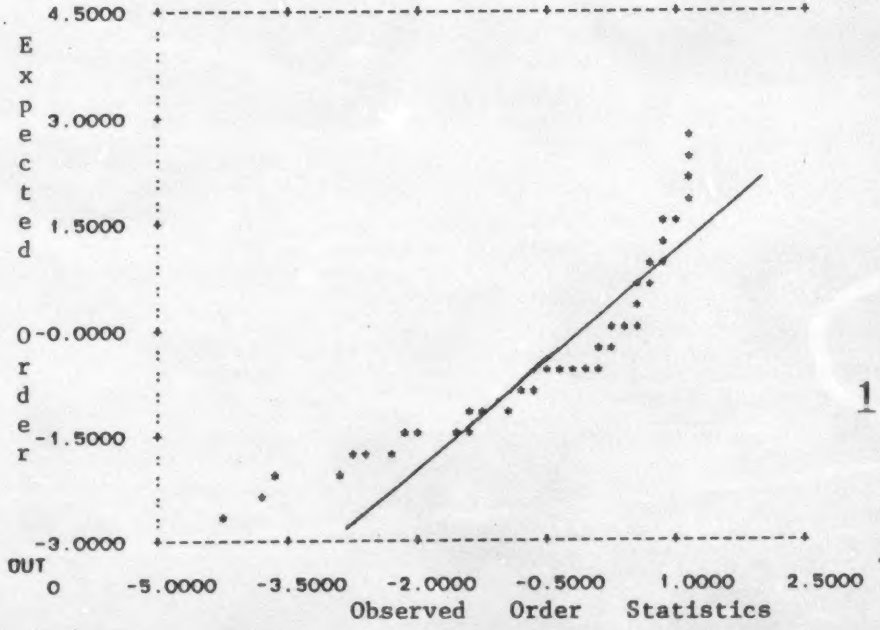
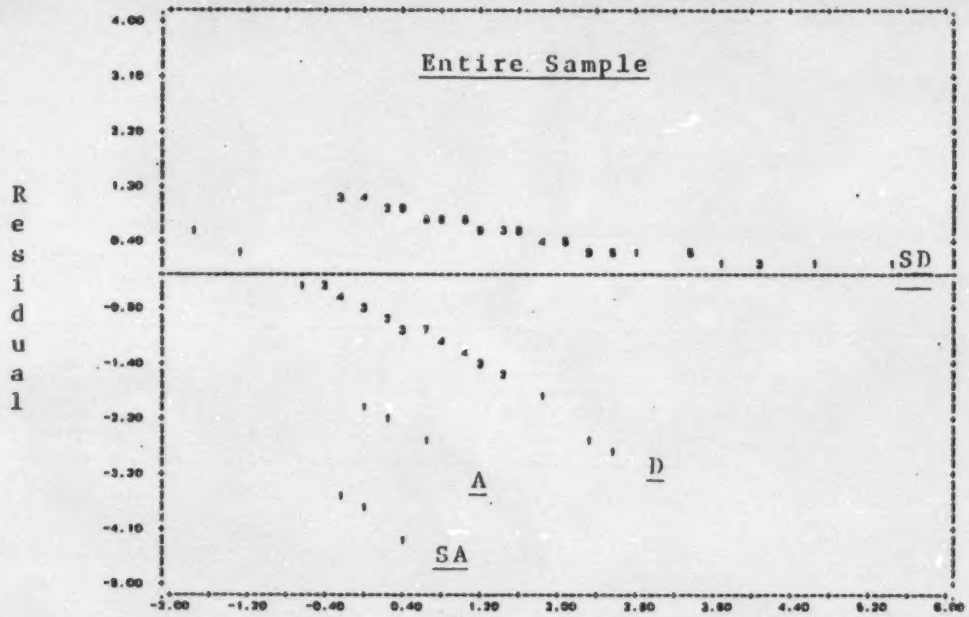
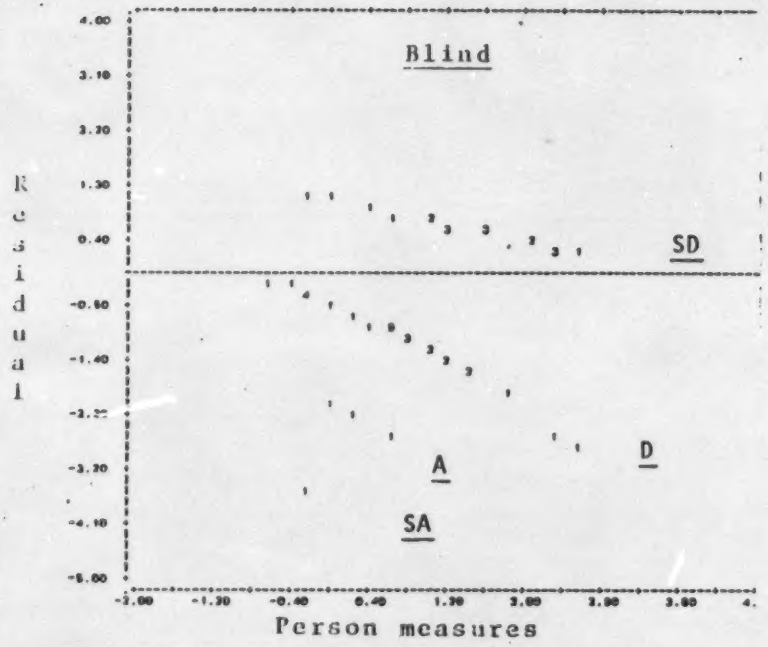


Figure 4

Residuals plotted against the attitude measures



1242: Blindness is a form of punishment for one's sins.



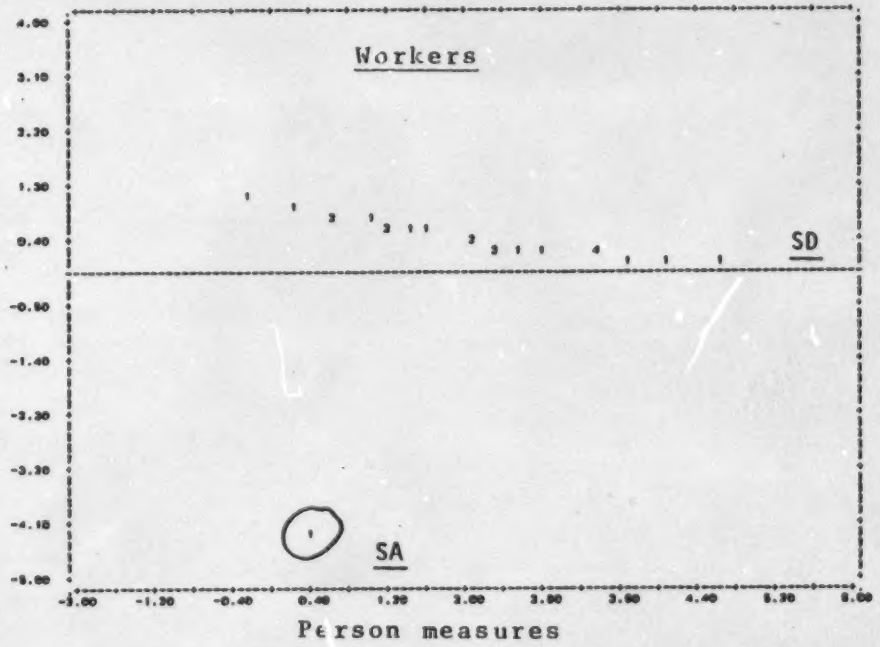


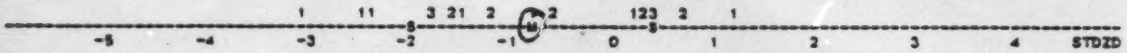
Figure 5

Residual Frequency Distributions

Figure 6

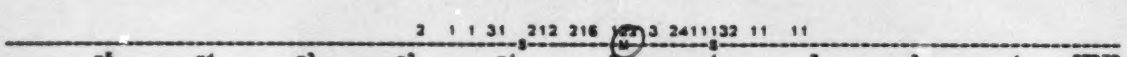
Work Excuse

REHAB WORKERS WHO ARE NOT BLIND



1101	M	S	N
	-0.78	1.20	23.

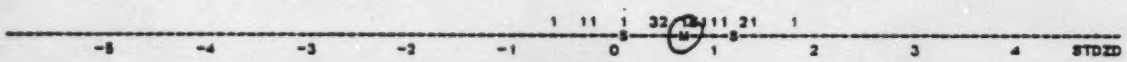
BLIND PERSONS WHO ARE NOT REHAB WORKERS



1101	M	S	N
	0.07	0.96	48.

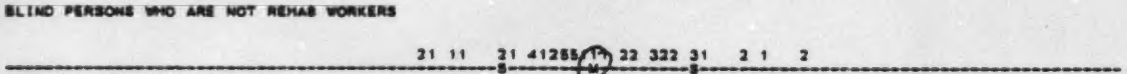
Supervisors

REHAB WORKERS WHO ARE NOT BLIND



1106	M	S	N
	0.67	0.87	23.

BLIND PERSONS WHO ARE NOT REHAB WORKERS

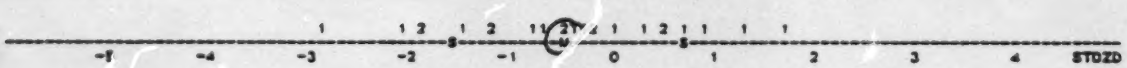


1106	M	S	N
	-0.10	0.93	48.

Sighted Uneasy

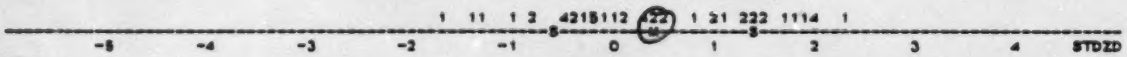
(hardest item)

REHAB WORKERS WHO ARE NOT BLIND



1206	M	S	N
	-0.45	1.15	23.

BLIND PERSONS WHO ARE NOT REHAB WORKERS



1206	M	S	N
	0.40	1.04	48.

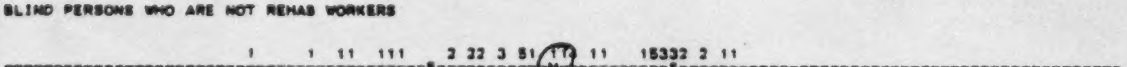
Punishment for one's sins

(easiest item)

REHAB WORKERS WHO ARE NOT BLIND

1242	M	S	N
	0.27	1.03	23.

BLIND PERSONS WHO ARE NOT REHAB WORKERS



1242	M	S	N
	-0.58	1.18	48.

Figure 7

Item residual means for the workers

plotted against the blind means